Kolmogorov's forward PIDE and forward transition rates in life insurance

Kristian Buchardt[1]
University of Copenhagen and PFA Pension

December 15, 2015

Abstract

We consider a doubly stochastic Markov chain, where the transition intensities are modelled as diffusion processes. Here we present a forward partial integro-differential equation for the transition probabilities. This is a generalisation of Kolmogorov's forward differential equation. In this setup, we define forward transition rates, generalising the concept of forward rates, e.g. the forward mortality rate. These models are applicable in e.g. life insurance mathematics, which is treated in the paper. The results presented follow from the general forward partial integro-differential equation for stochastic processes, of which the Fokker-Planck differential equation and Kolmogorov's forward differential equation are the two most known special cases. We end the paper by considering the semi-Markov case, which can also be considered a special case of a general forward partial integro-differential equation.

Keywords: Kolmogorov's differential equation; doubly stochastic Markov model; stochastic transition rate; forward transition rate; life insurance; semi-Markov

1 Introduction

In this paper we present Kolmogorov's forward partial integro-differential equation for a jump-diffusion. We study the doubly stochastic setup in particular, in which a finite state jump process has transition rates which are driven by a diffusion process. We show how this can be applied in life insurance. This motivates the definition of so-called forward transition rates, which can be written as a conditional expectation of the stochastic transition rates. With these, the usual results from the classic life insurance

[1]Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark and (permanent address) PFA Pension, Sundkrogsgade 4, DK-2100 Copenhagen Ø, Denmark, e-mail: `kristian@buchardt.net`

setup hold, where the forward transition rates replace the usual transition rates, and in particular the traditional version of Kolmogorov's forward differential equation holds. We end the paper by studying the semi-Markov setup, where the transition rates are not driven by a diffusion, but simply depend on the duration in the current state. Here Kolmogorov's forward integro-differential equation is well known, and we show how it is a special case of the general forward partial integro-differential equation.

We begin this paper by presenting a forward partial integro-differential equation for the transition probabilities of a general jump-diffusion, which is inspired by Chapter 3.4 in [9]. A forward partial differential equation exists for the case of a continuous diffusion, and this is known as the Fokker-Planck differential equation. For the case of a pure jump process, this is also known, and is called Kolmogorov's forward differential equation or the master equation. The more general result presented here we refer to as *Kolmogorov's forward partial integro-differential equation*, and this sets the basis for the rest of the results.

The main case studied in this paper is the doubly stochastic Markov chain setup. This consists of a stochastic process $\mathbf{Z} = (Z(t))_{t\geq 0}$ taking values in a finite state space $\mathcal{J} = \{0, 1, \ldots, J\}$. The transition rates of $\mathbf{Z}$ are driven by a stochastic diffusion process $\mathbf{X} = (X(t))_{t\geq 0}$, and $(\mathbf{Z}, \mathbf{X})$ is a Markov chain. We refer to $\mathbf{Z}$ as a doubly stochastic Markov chain. A main result of this paper is that we present Kolmogorov's forward differential equation for this setup, as a special case of the general equation. It is a partial integro-differential equation, and can be considered a generalisation of the traditional version of Kolmogorov's forward differential equation. One possible application of this setup is in life insurance, where $\mathbf{Z}$ models the state of a policyholder, e.g. alive, disabled, dead etc. The transition rates between the different states are not known in the future, and are then modelled as a diffusion process. The transition probabilities of $\mathbf{Z}$ are essential for calculating the expected future cash flow for such a life insurance policy, and they can be found using Kolmogorov's forward partial integro-differential equation. It is well known that a backward partial differential equation exists, but for calculating e.g. life insurance cash flows, a forward differential equation is more efficient in practice. For more on the doubly stochastic setup in life insurance, see [17], [6], [1], and in credit risk, see e.g. [13]. A special case of the doubly stochastic setup is the classic life insurance setup, where the transition rates of $\mathbf{Z}$ are deterministic. In that case Kolmogorov's forward differential equation is well known, and for more on this setup in life insurance, see e.g. [12] or [18], and for cash flows, see [2].

For the doubly stochastic Markov chain we define so-called forward transition rates, which is one of the main contributions of this paper. We show that these can be represented as a conditional expectation of the (stochastic) transition rates. With the forward transition rates, Kolmogorov's forward differential equation from the classic life insurance setup holds, which can be used to calculate the transition probabilities of $\mathbf{Z}$ directly.

This generalises the concept of forward rates for doubly stochastic Markov chains, where the best known example is the forward mortality rate, treated in e.g. [14], [6], [3] and [7]. The concept also relates to the dependent forward rates from [1], and these relations are studied in the current paper. For more on forward rates, see [5] where they investigate classes of models where the same forward rates can be used, and [19], which gives a critical overview of the concept of forward rates in life- and pension insurance. In the classic life insurance setup where the transition rates are deterministic so $\mathbf{Z}$ is a Markov chain, the forward transition rates simplify and equal the transition rates. To the authors knowledge, a general definition for forward transition rates for doubly stochastic Markov chains has not been presented before.

Another example of a stochastic process in life insurance is the so-called semi-Markov setup. Here, $\mathbf{Z}$ is a stochastic process on a finite state space, and we define the duration of $\mathbf{Z}$ in the current state as $\mathbf{U} = (U(t))_{t\geq 0}$. Then, if $(\mathbf{Z}, \mathbf{U})$ is a Markov process, $\mathbf{Z}$ is a semi-Markov process. In this case, the transition rates may depend on the duration $\mathbf{U}$. It is already known that a version of Kolmogorov's forward differential equation exists for this case. For more on semi-Markov models see e.g. [2], [4] or [10], and see also [8] where they refer to the model as a piecewise-deterministic Markov process. In this paper, we show how Kolmogorov's forward partial integro-differential equation specialises in the semi-Markov setup to Kolmogorov's forward integro-differential equation.

The structure of the paper is as follows. In Section 2, we consider a jump-diffusion Markov process and find a forward partial integro-differential equation for the transition probabilities, Kolmogorov's forward partial integro-differential equation. In Section 3, we consider the doubly stochastic setup where the transition rates of $\mathbf{Z}$ is driven by a continuous diffusion process $\mathbf{X}$ and present Kolmogorov's forward partial integro-differential equation for this case. Following this, we define the forward transition rates in Section 4, and relate them to existing literature. In Section 5, we show how to apply the doubly stochastic setup and the forward transition rates in life insurance. Finally, in Section 6, we relate Kolmogorov's forward partial integro-differential equation to the semi-Markov setup, and see that we obtain the same integro-differential equation as in [2].

2 The forward partial integro-differential equation

Let $\mathbf{X}$ be a d-dimensional stochastic jump-diffusion in a state space $\mathcal{X} \subset \mathbb{R}^d$, defined on a probability space $(\Omega, \mathbb{F}, \mathrm{P})$. We assume that $\mathbf{X}$ is a solution to the stochastic differential equation,

$$\mathrm{d}X(t) = \beta(t, X(t))\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t) + \mathrm{d}J(t), \tag{2.1}$$

for all $t \in [0,T]$. Throughout the article, T is some finite time-horizon. Let $\beta : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^d$, $\beta \in \mathcal{C}^{1,1}$ and $\sigma : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$, $\sigma \in \mathcal{C}^{1,2}$, where $\mathcal{C}^{n,m}$ is the set of functions that are n times continuously differentiable in the first argument, and m times continuously differentiable in the second argument. Let $\mathbf{W}$ be a d-dimensional standard Brownian motion, and define $\rho(t,x) = \sigma(t,x)\sigma(t,x)^\top$. The jump part of $\mathbf{X}$ is $\mathbf{J}$, which has a jump distribution that only depend on the current time and state of $\mathbf{X}$. We assume existence of a jump intensity measure, which is denoted $\mu : \mathbb{R}^d \times \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^+$. Further assume that $\mu(\mathbb{R}^d; t,x)$ is bounded for all t,x, and that the compensated process

$$t \mapsto J(t) - \int_0^t \int (y - X(s-))\,\mu(\mathrm{d}y; s, X(s-))\,\mathrm{d}s \tag{2.2}$$

is a martingale. Thus, $\mu(\mathrm{d}y;t,x)$ is the measure of the jump destinations of $\mathbf{X}$, and not jump sizes.

In (2.2) and in general, when we omit integration bounds we integrate over the whole domain of the integrand. We use the notation $f(x-) = \lim_{y\nearrow x} f(y)$, and use τ_A as the counting measure on the set A.

Last, we define $\mathbf{N}$ as the process counting the number of jumps,

$$N(t) = \#\{s \in (0,t] \mid J(s-) \neq J(s)\},$$

and it follows that $\mathbf{N}$ has compensator $\int_0^t \mu(\mathbb{R}^d; s, X(s-))\mathrm{d}s$.

The transition probability of $\mathbf{X}$ is denoted P, that is, the distribution of $X(t)$ given the value at some earlier time point. For $t \geq t'$ and a Borel set $A \subset \mathbb{R}^d$ we write

$$\mathrm{P}(X(t) \in A \mid X(t') = x') = P(t, A; t', x') = \int_A P(t, \mathrm{d}x; t', x').$$

The transition probability can be described by a forward partial integro-differential equation (PIDE). This result sets the basis for this article.

Theorem 2.1. (Kolmogorov's forward PIDE) *Assume for $i = 1, \ldots, d$ that*

$$\int_0^t \beta_i(s, X(s)) f(X(s))\,\mathrm{d}W(s)$$

is a martingale, for all $f \in \mathcal{C}^1$ with compact support. Assume there exists a set $\tilde{\mathcal{X}} \subset \mathcal{X}$ such that $\frac{\partial}{\partial t}P(t,\mathrm{d}x;t',x')$, $\frac{\partial}{\partial x_i}P(t,\mathrm{d}x;t',x')$ and $\frac{\partial^2}{\partial x_i \partial x_j}P(t,\mathrm{d}x;t',x')$ exist for all $i,j = 1,\ldots,d$, all $x \in \tilde{\mathcal{X}}$ and $t \in (t', T]$. Then the transition probability P of $\mathbf{X}$ satisfies, for

$t \in (t', T]$, *the PIDE,*

$$\begin{aligned}\frac{\partial}{\partial t}P(t,A;t',x') = &-\sum_i \int_A \frac{\partial}{\partial x_i}\left(\beta_i(t,x)P(t,\mathrm{d}x;t',x')\right) \\ &+\frac{1}{2}\sum_{i,j}\int_A \frac{\partial^2}{\partial x_i \partial x_j}\left(\rho_{ij}(t,x)P(t,\mathrm{d}x;t',x')\right) \\ &+\int_{A^{\complement}}\left(\int_A \mu(\mathrm{d}x;t,y)\right)P(t,\mathrm{d}y;t',x') \\ &-\int_A\left(\int_{A^{\complement}}\mu(\mathrm{d}y;t,x)\right)P(t,\mathrm{d}x;t',x'),\end{aligned} \tag{2.3}$$

for any compact Borel set $A \subset \tilde{\mathcal{X}}$.

The notation $A^{\complement}$ is the complement of the set A.

Remark 2.2. The transition probability trivially satisfies the boundary condition

$$P(t',A;t',x) = 1_A(x).$$

$\diamond$

This differential equation does not seem to have any agreed name in the literature. If $\mathbf{X}$ is a continuous diffusion process, it is called the Fokker-Planck equation, and if $\mathbf{X}$ is a pure jump process it is often called Kolmogorov's forward differential equation, and also the master equation. Since Kolmogorov's forward differential equation seems to be the adopted term and the actuarial and financial literature, and we consider this a generalisation, we refer to it as *Kolmogorov's forward PIDE.* For more about Kolmogorov's differential equations and applications in life insurance, see e.g. [18]. For more on the Fokker-Planck and the master equation, see e.g. [9].

The following proof is inspired by the calculations in [9], and the main difference from there is that in [9] it is implicitly assumed that the transition probability $P(t,\mathrm{d}x;t',x')$ has a density with respect to the Lebesgue measure. In this paper, we do not generally assume existence of a density with respect to the Lebesgue measure. Further, the proof in [9] is not based on a direct application of Itô's lemma.

Proof. Let $f \in \mathcal{C}^2$, and assume that f has support on a compact set $A \subset \tilde{\mathcal{X}}$. An application of Itô's lemma yields,

$$
\begin{aligned}
&f(X(t)) - f(X(t')) \\
&\quad = \int_{t'}^{t} \left\{ \sum_i \beta_i(s, X(s)) \frac{\partial f(X(s))}{\partial x_i} + \frac{1}{2} \sum_{i,j} \rho_{ij}(s, X(s)) \frac{\partial^2 f(X(s))}{\partial x_i \partial x_j} \right\} \mathrm{d}s \\
&\qquad + \int_{t'}^{t} \int \left(f(y) - f(X(s-)) \right) \mu(\mathrm{d}y; s, X(s-)) \mathrm{d}s \\
&\qquad + \int_{t'}^{t} \sum_i \beta_i(s, X(s)) \frac{\partial f(X(s))}{\partial x_i} \mathrm{d}W(s) + \tilde{J}(t) - \tilde{J}(t'),
\end{aligned}
\tag{2.4}
$$

where $\tilde{J}(t) = \int_0^t \left(f(X(s)) - f(X(s-)) \right) \mathrm{d}N(s) - \int_0^t \int \left(f(y) - f(X(s-)) \right) \mu(\mathrm{d}y; s, X(s-)) \mathrm{d}s$. By the assumption (2.2) we infer that $\tilde{J}(t)$ is a martingale. We also have,

$$
\int_A f(x) P(t, \mathrm{d}x; t', x') - f(x') = \mathrm{E}\left[f(X(t)) - f(X(t')) \middle| X(t') = x' \right].
$$

By insertion of (2.4), and using that the last line of (2.4) is a martingale, we obtain

$$
\begin{aligned}
&\int_A f(x) P(t, \mathrm{d}x; t', x') - f(x') \\
&\quad = \mathrm{E}\Bigg[\int_{t'}^{t} \Bigg\{ \sum_i \beta_i(s, X(s)) \frac{\partial f(X(s))}{\partial x_i} + \frac{1}{2} \sum_{i,j} \rho_{ij}(s, X(s)) \frac{\partial^2 f(X(s))}{\partial x_i \partial x_j} \Bigg\} \mathrm{d}s \\
&\qquad\quad + \int_{t'}^{t} \int \left(f(y) - f(X(s-)) \right) \mu(\mathrm{d}y; s, X(s-)) \mathrm{d}s \Bigg| X(t') = x' \Bigg] \\
&\quad = \int_{t'}^{t} \int_A \Bigg\{ \sum_i \beta_i(s, x) \frac{\partial f(x)}{\partial x_i} + \frac{1}{2} \sum_{i,j} \rho_{ij}(s, x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \Bigg\} P(s, \mathrm{d}x; t', x') \mathrm{d}s \\
&\qquad\quad + \int_{t'}^{t} \int \int \left(f(y) - f(x) \right) \mu(\mathrm{d}y; s, x) P(s, \mathrm{d}x; t', x') \mathrm{d}s.
\end{aligned}
$$

We differentiate with respect to t, and then apply partial integration to the two first terms,

$$
\begin{aligned}
\frac{\partial}{\partial t} \int_A f(x) P(t, \mathrm{d}x; t', x') = & \sum_i \int_A \beta_i(t, x) \frac{\partial f(x)}{\partial x_i} P(t, \mathrm{d}x; t', x') \\
& + \frac{1}{2} \sum_{i,j} \int_A \rho_{ij}(t, x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} P(t, \mathrm{d}x; t', x') \\
& + \int \int \left(f(y) - f(x) \right) \mu(\mathrm{d}y; t, x) P(t, \mathrm{d}x; t', x')
\end{aligned}
$$

$$
\begin{aligned}
&= -\sum_i \int_A f(x) \frac{\partial}{\partial x_i} \left(\beta_i(t,x) P(t, \mathrm{d}x; t', x')\right) \\
&\quad + \frac{1}{2} \sum_{i,j} \int_A f(x) \frac{\partial^2}{\partial x_i \partial x_j} \left(\rho_{ij}(t,x) P(t, \mathrm{d}x; t', x')\right) \\
&\quad + \int \int \left(f(y) - f(x)\right) \mu(\mathrm{d}y; t, x) P(t, \mathrm{d}x; t', x').
\end{aligned}
$$

The boundary terms from the partial integration vanishes due to the fact that $f(x) = 0$ and $\frac{\partial}{\partial x_i} f(x) = 0$ on the boundary of A, since f is $\mathcal{C}^2$.

Let $f_n \in \mathcal{C}^2$ be a series of uniformly bounded functions with compact support in $\tilde{\mathcal{X}}$, such that $f_n \to 1_A$ for $n \to \infty$. This yields the result. $\square$

In (2.3) and in general, we differentiate the integrator of an integral, e.g. in $\int_A \frac{\partial}{\partial x_i} (\beta_i(t,x) P(t, \mathrm{d}x; t', x'))$. This notation is adopted to be able to write a general expression, where we do not need to assume a density for P with respect to another measure. For a specific model for $\mathbf{X}$, one would typically determine a density and reference measure so P can be written in terms of these. This reference measure could be the Lebesgue or the counting measure on some set, or a mixture. This is for example the case in Corollary 3.2.

If $\mu(\mathbb{R}^d; t, x) = 0$ for all t, x, the process $\mathbf{X}$ is continuous. In that case, (2.3) reduces to an integral version of the well-known Fokker-Planck equation.

$$
\begin{aligned}
\frac{\partial}{\partial t} P(t, A; t', x') = &-\sum_i \int_A \frac{\partial}{\partial x_i} \left(\beta_i(t,x) P(t, \mathrm{d}x; t', x')\right) \\
&+ \frac{1}{2} \sum_{i,j} \int_A \frac{\partial^2}{\partial x_i \partial x_j} \left(\rho_{ij}(t,x) P(t, \mathrm{d}x; t', x')\right).
\end{aligned}
$$

In the opposite case, where $\beta(t,x) = 0$ and $\sigma(t,x) = 0$ for all t, x, the process $\mathbf{X}$ is a pure jump process. In that case, (2.3) reduces to an integral version of Kolmogorov's forward differential equation, also known as the master equation,

$$
\begin{aligned}
\frac{\partial}{\partial t} P(t, A; t', x') = &\int_{A^{\complement}} \left(\int_A \mu(\mathrm{d}x; t, y)\right) P(t, \mathrm{d}y; t', x') \\
&- \int_A \left(\int_{A^{\complement}} \mu(\mathrm{d}y; t, x)\right) P(t, \mathrm{d}x; t', x'),
\end{aligned}
$$

Assuming that a density exists with respect to some measure, one can differentiate and obtain the Fokker-Planck respectively Kolmogorov's forward differential equation.

In particular the jump part of (2.3) is easily interpreted. The positive term leads to an increasing probability, and it is the probability of being somewhere in the complement of

A and making a jump inside A. The negative term leads to a decreasing probability, and it is the probability of being somewhere in A, and making a jump to the complement of A. Jumps solely inside A or $A^{\complement}$ does not affect the probability.

3 The doubly stochastic Markov chain setup

In this section we consider the doubly stochastic Markov chain setup, which can be considered a special case of the stochastic process $\mathbf{X}$ from (2.1). We let a set of stochastic rates be modelled as a continuous diffusion process. Conditional on these, we create a Markov chain with these transition rates. For simplicity we restrict to the case of continuous transition rates.

Let a finite state space $\mathcal{J} = \{0, \ldots, J\}$ be given. For each $k, \ell \in \mathcal{J}$, where $k \neq \ell$, we associate a transition rate $t \mapsto \mu_{k\ell}(t, X(t))$. Here, $\mathbf{X}$ is a d-dimensional diffusion process satisfying the stochastic differential equation

$$\mathrm{d}X(t) = \beta(t, X(t))\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t),$$

where $\mathbf{W}$ is a d-dimensional diffusion process and β and σ are as in Section 2. Thus, the transition rates $\mu_{k\ell}$ are stochastic.

Define now the stochastic process $\mathbf{Z}$ on $\mathcal{J}$, and assume it is cádlág. Conditional on $\mathbf{X}$, we let $\mathbf{Z}$ be a Markov chain, with transition rates $\mu_{k\ell}(t, X(t))$. Define the filtrations generated by $\mathbf{X}$ respectively $\mathbf{Z}$ as $\mathcal{F}^{\mathbf{X}}(t) = \sigma(X(s)|s \leq t)$ and $\mathcal{F}^{\mathbf{Z}}(t) = \sigma(Z(s)|s \leq t)$. Define also the larger filtration $\mathcal{F}(t) = \mathcal{F}^{\mathbf{X}}(t) \vee \mathcal{F}^{\mathbf{Z}}(t)$. Let $\mathbf{N}_{k\ell}$ be a counting process that counts the number of jumps from state k to state ℓ,

$$N_{k\ell}(t) = \#\{s \leq t | Z(s-) = k, Z(s) = \ell\}. \tag{3.1}$$

The assumption that $\mathbf{Z}$ is a Markov chain with stochastic transition rates $\mu_{k\ell}(t, X(t))$ means, that conditional on $\mathbf{X}$, the compensated process

$$N_{k\ell}(t) - \int_0^t 1_{(Z(s-)=k)} \mu_{k\ell}(s, X(s))\, \mathrm{d}s \tag{3.2}$$

is a martingale. That is, it is a martingale with respect to the filtration $\mathcal{F}^{\mathbf{X}}(T) \vee \mathcal{F}^{\mathbf{Z}}(t)$. In particular, it is straightforward to verify that (3.2) is also a martingale unconditionally on $\mathbf{X}$, that is, with respect to the filtration $\mathcal{F}(t)$.

We are interested in finding transition probabilities for the doubly stochastic Markov chain $Z(t)$,

$$p_{k\ell}(t; t', x) = P(Z(t) = \ell \mid Z(t') = k, X(t') = x). \tag{3.3}$$

These are dependent on $X(t')$; we think of time t' as *now*, and thus $X(t')$ is known. We find this transition probability by first finding the transition probability for the combined process $(Z(t), X(t))$. Thus, let $P(t, k, A; t', k', x')$ denote the transition probability of $(Z(t), X(t))$ conditional on $(Z(t'), X(t')) = (k', x')$. Considering the $(d+1)$-dimensional process $(Z(t), X(t))$ as a special case of (2.1), we obtain the following result.

Corollary 3.1. (Kolmogorov's forward PIDE for the doubly stochastic setup) *Assume that*

$$\int_0^t \beta_i(s, X(s)) f(X(s)) \, \mathrm{d}W(s)$$

is a martingale for all $f \in \mathcal{C}^1$ with compact support, and that $\frac{\partial}{\partial t} P(t, k, \mathrm{d}x; t', k', x')$, $\frac{\partial}{\partial x_i} P(t, k, \mathrm{d}x; t', k', x')$ and $\frac{\partial^2}{\partial x_i \partial x_j} P(t, k, \mathrm{d}x; t', k', x')$ exist. Then the transition probability $P(t, k, A; t', k', x')$ for $k \in \mathcal{J}$ and a compact Borel-set $A \subset \mathbb{R}^d$ satisfy the forward PIDE

$$\begin{aligned}
\frac{\partial}{\partial t} P(t, k, A; t', k', x') = &- \sum_i \int_A \frac{\partial}{\partial x_i} \left(\beta_i(t, x) P(t, k, \mathrm{d}x; t', k', x') \right) \\
&+ \frac{1}{2} \sum_{i,j} \int_A \frac{\partial^2}{\partial x_i \partial x_j} \left(\rho_{ij}(t, x) P(t, k, \mathrm{d}x; t', k', x') \right) \\
&+ \sum_{\ell : \ell \neq k} \int_A \mu_{\ell k}(t, x) P(t, \ell, \mathrm{d}x; t', k', x') \\
&- \int_A \sum_{\ell : \ell \neq k} \mu_{k \ell}(t, x) P(t, k, \mathrm{d}x; t', k', x'),
\end{aligned} \tag{3.4}$$

subject to the boundary condition $P(t', k, A; t', k', x) = 1_{(k=k')} 1_A(x)$.

Proof. We have the dynamics

$$\mathrm{d} \begin{bmatrix} Z(t) \\ X(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \beta(t, X(t)) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 & 0 \\ 0 & \sigma(t, X(t)) \end{bmatrix} \mathrm{d}\widetilde{W}(t) + \begin{bmatrix} \mathrm{d}J(t) \\ 0 \end{bmatrix},$$

for $\widetilde{W}(t) = (\hat{W}(t), W(t))$, where $\hat{W}(t)$ is an adapted standard Brownian motion. In particular, the jump measure $\tilde{\mu}(\mathrm{d}(\ell, y); t, (k, x))$ of this process is,

$$\tilde{\mu}(\mathrm{d}(\ell, y); t, (k, x)) = \mu_{k\ell}(t, x) \cdot \mathrm{d} \left(\tau_{\mathcal{J} \setminus \{k\}}(\ell) \otimes \tau_{\{x\}}(y) \right).$$

Now the result follows from Theorem 2.1 applied to the process $(Z(t), X(t))$. $\square$

Corollary 3.2. *Assume Corollary 3.1 holds and that a density with respect to the Lebesgue measure exists,*

$$P(t, k, \mathrm{d}x; t', k', x') = p(t, k, x; t', k', x') \, \mathrm{d}x. \tag{3.5}$$

Then the density satisfies, for $t > t'$, the PDE

$$\begin{aligned}
\frac{\partial}{\partial t} p(t,k,x;t',k',x') = &-\sum_{i=1}^{d} \frac{\partial}{\partial x_i} \left(\beta_k(t,x) p(t,k,x;t',k',x')\right) \\
&+ \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} \left(\rho_{ij}(t,x) p(t,k,x;t',k',x')\right) \\
&+ \sum_{\ell;\ell\neq k} \mu_{\ell k}(t,x) p(t,\ell,x;t',k',x') \\
&- \sum_{\ell;\ell\neq k} \mu_{k\ell}(t,x) p(t,k,x;t',k',x'),
\end{aligned} \tag{3.6}$$

subject to the boundary condition $p(t',k,x;t',k',x') = 1_{(k=k')} 1_{(x=x')}$.

Assuming we have solved the PIDE (3.4) or the PDE (3.6), we have the transition probability for the process $(Z(t), X(t))$. If we are interested in the transition probabilities of $Z(t)$ only, (3.3), we can integrate over the underlying state $X(t)$,

$$p_{k'k}(t;t',x') = P(t,k,\mathbb{R}^d;t',k',x') = \int p(t,k,x;t',k',x') \, \mathrm{d}x. \tag{3.7}$$

Corollary 3.2 is a generalisation of Kolmogorov's forward differential equation. If the transition rates are deterministic, which without loss of generality can be characterised as $\mathbf{X}$ being constant, we have that $\beta(t,x) = 0$ and $\sigma(t,x) = 0$. In that case, (3.6) simplifies to Kolmogorov's forward differential equation,

$$\frac{\partial}{\partial t} p(t,k;t',k') = \sum_{\ell \in \mathcal{J}\setminus\{k\}} \left(\mu_{\ell k}(t) p(t,\ell;t',k') - \mu_{k\ell}(t) p(t,k;t',k')\right).$$

Here we removed x, y from the notation in $p(s,j,y;t,i,x)$ and $\mu_{k\ell}(t,x)$.

3.1 Backward partial differential equation

It is well known that a backward PDE exists for the transition probability; since

$$t \mapsto \mathrm{E}\left[1_{(Z(s)=k)} \middle| \mathcal{F}(t)\right] = \sum_{\ell} 1_{(Z(t)=\ell)} p_{\ell k}(s;t,X(t)),$$

is a martingale, we can apply Itô's lemma and set the drift equal to zero. This yields a PDE in t, x, ℓ for $p_{\ell k}(s;t,x)$, together with the boundary conditions $p_{\ell k}(s;s,x) = 1_{(\ell=k)}$.

Proposition 3.3. (Kolmogorov's backward PDE) *For $\ell \in \mathcal{J}$, the transition probabilities $p_{\ell k}(s;t,x)$ satisfy the backward PDE*

$$\begin{aligned}\frac{\partial}{\partial t} p_{\ell k}(s;t,x) = &- \sum_{i=1}^{d} \beta_i(t,x) \frac{\partial}{\partial x_i} p_{\ell k}(s;t,x) - \frac{1}{2} \sum_{i,j=1}^{d} \rho_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} p_{\ell k}(s;t,x) \\ &- \sum_{m;m\neq \ell} \mu_{\ell m}(t,x) \left(p_{mk}(s;t,x) - p_{\ell k}(s;t,x) \right),\end{aligned} \tag{3.8}$$

subject to the boundary condition $p_{\ell k}(s;s,x) = 1_{(\ell = k)}$.

If the transition rates are deterministic, which can be modelled by setting $\beta(t,x) = 0$ and $\rho(t,x) = 0$, the first two terms of (3.8) disappear, and we are left with Kolmogorov's well known backward differential equation.

With the results presented, we have two ways of calculating $p_{\ell k}(s;t,x)$. Either, we solve the backward PDE in Proposition 3.3, or we solve the forward PDE from Corollary 3.2 and integrate over the transition rates as in (3.7). In Section 5 about life insurance cash flows, we see that we need to find $p_{\ell k}(s;t,x)$ for fixed ℓ, t, x and varying k, s. In that case, the forward PDE seems preferable, as we only need to solve it once. If we use the backward PDE, we need to solve it for every fixed pair of k, s.

4 Forward transition rates

In the doubly stochastic setup, a particular quantity of interest is the transition probabilities of $\mathbf{Z}$, from (3.3),

$$p_{k'k}(t;t',x') = \int p(t,k,x;t',k',x')\,\mathrm{d}x.$$

They are used if we are only interested in $\mathbf{Z}$ and not the underlying stochastic process driving the transition rates, e.g. for valuation of life insurance contracts, see Section 5. By the definition of so-called *forward transition rates*, we are able to present a differential equation for this transition probability directly. The differential equation is identical to Kolmogorov's forward differential equation which is used when the transition rates are deterministic. We define the forward transition rate as the instantaneous probability of a transition from k to ℓ at time t. This is dependent on the state of $(\mathbf{Z}, \mathbf{X})$ at the present time t'.

Definition 4.1. *Given it exists, the* forward transition rate *$f_{k\ell}(t;t')$ at time t' is defined as*

$$f_{k\ell}(t;t') = \lim_{h \searrow 0} \frac{1}{h} \mathrm{P}\left(Z(t+h) = \ell \,\middle|\, Z(t) = k, (Z,X)(t') \right).$$

Remark 4.2. The forward transition rate at time t' is $(Z,X)(t')$-measurable. We may condition on the exact values of $(Z,X)(t')$ and write

$$f_{k\ell}(t;t',k',x') = \lim_{h\searrow 0}\frac{1}{h}\,\mathrm{P}\left(Z(t+h)=\ell\,\middle|\,Z(t)=k,(Z,X)(t')=(k',x')\right).$$

Then $f_{k\ell}(t;t',Z(t'),X(t')) = f_{k\ell}(t,t')$. $\diamond$

The classic setup, where $\mathbf{Z}$ is itself Markov and independent of $\mathbf{X}$, is a special case of the doubly stochastic setup considered here. In that case, by independence, we see that $\mathbf{X}$ disappears from the conditioning in Definition 4.1, and that $Z(t')$ disappears as well, due to the Markov property. This is exactly the definition of the transition rate in the classic setup, and Definition 4.1 is a generalisation of the transition rates to the doubly stochastic setup.

Assuming that the transition rates exist, we can represent the forward transition rates as an expectation of the transition rates.

Lemma 4.3. *If the transition rates $\mu_{k\ell}(t,X(t))$ exist, then*

$$f_{k\ell}(t;t',k',x') = \mathrm{E}\left[\mu_{k\ell}(t,X(t))\middle|\,Z(t)=k,(Z,X)(t')=(k',x')\right].$$

Proof. From Definition 4.1 and Remark 4.2, we condition on $\mathbf{X}$ and interchange expectation and limits,

$$\begin{aligned}
f_{k\ell}(t;t') &= \lim_{h\searrow 0}\frac{1}{h}\,\mathrm{P}\left(Z(t+h)=\ell\,\middle|\,Z(t)=k,(Z,X)(t')\right)\\
&= \mathrm{E}\left[\lim_{h\searrow 0}\frac{1}{h}\,\mathrm{P}\left(Z(t+h)=\ell\,\middle|\,\mathbf{X},Z(t)=k,(Z,X)(t')\right)\middle|\,Z(t)=k,(Z,X)(t')\right]\\
&= \mathrm{E}\left[\lim_{h\searrow 0}\frac{1}{h}\,\mathrm{P}\left(Z(t+h)=\ell\,|\mathbf{X},Z(t)=k\right)\middle|\,Z(t)=k,(Z,X)(t')\right]\\
&= \mathrm{E}\left[\mu_{k\ell}(t,X(t))\middle|\,Z(t)=k,(Z,X)(t')\right],
\end{aligned}$$

where we at third line used that conditional on $\mathbf{X}$, $\mathbf{Z}$ is a Markov chain, and at the fourth line we used the definition of the transition rate of a Markov chain. $\square$

The transition rates $\mu_{k\ell}(t,X(t))$ determine the behaviour of $Z(t)$ in the next infinitesimal timespan, and this is dependent on $X(t)$. If we at present time t' are interested in the behaviour of $Z(t)$ at the future time point t, we can average out the dependence on $X(t)$, and then we obtain the forward transition rate $f_{k\ell}(t;t')$. These forward rates determine the expected behaviour of $Z(t)$ conditional on the present state at time t', and with this we can write Kolmogorov's forward differential equation on the usual form, with the forward transition rates in place of the transition rates.

Theorem 4.4. *Assume the forward transition rates of Definition 4.1 exist. Then the transition probabilities $p_{k'k}(t;t',x')$ for $\mathbf{Z}$ satisfy the differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{k'k}(t;t',x') = \sum_{\ell;\ell\neq k} f_{\ell k}(t;t',k',x')p_{k'\ell}(t;t',x') - \sum_{\ell;\ell\neq k} f_{k\ell}(t;t',k',x')p_{k'k}(t;t',x'), \tag{4.1}$$

with boundary conditions $p_{k'k}(t';t',x') = 1_{\{k'=k\}}$.

Proof. We apply the Chapman-Kolmogorov equation and rearrange,

$$\begin{aligned}
&p_{k'k}(t+h;t',x') \\
&= \sum_{\ell} p_{k'\ell}(t;t',x')\,\mathrm{P}\left(Z(t+h)=k \,\middle|\, Z(t)=\ell,(Z,X)(t')=(k',x')\right) \\
&= \sum_{\ell;\ell\neq k} p_{k'\ell}(t;t',x')\,\mathrm{P}\left(Z(t+h)=k \,\middle|\, Z(t)=\ell,(Z,X)(t')=(k',x')\right) \\
&\quad + p_{k'k}(t;t',x')\left(1-\sum_{m;m\neq k}\mathrm{P}\left(Z(t+h)=m \,\middle|\, Z(t)=k,(Z,X)(t')=(k',x')\right)\right).
\end{aligned}$$

By multiplication with $\frac{1}{h}$,

$$\begin{aligned}
&\frac{p_{k'k}(t+h;t',x')-p_{k'k}(t;t',x')}{h} \\
&= \sum_{\ell;\ell\neq k} p_{k'\ell}(t;t',x')\frac{1}{h}\,\mathrm{P}\left(Z(t+h)=k \,\middle|\, Z(t)=\ell,(Z,X)(t')=(k',x')\right) \\
&\quad - p_{k'k}(t;t',x')\sum_{m;m\neq k}\frac{1}{h}\,\mathrm{P}\left(Z(t+h)=m \,\middle|\, Z(t)=k,(Z,X)(t')=(k',x')\right).
\end{aligned}$$

By letting $h \searrow 0$ and applying Definition 4.1, the result is obtained. $\square$

We can interpret the forward transition rate $f_{k\ell}(t;t',k',x')$ and the differential equation (4.1). At present time t', use the forward transition rates $f_{k\ell}(t;t',k',x')$ to define a new Markov chain $\widetilde{\mathbf{Z}}^{(t',k',x')}$, with $f_{k\ell}(t;t',k',x')$ as its deterministic transition rates, and with starting value $\widetilde{Z}^{(t',k',x')}(t') = k'$. With this construction, we know from Theorem 4.4 that $Z(t)$ and $\widetilde{Z}^{(t',k',x')}(t)$ equals in distribution in the conditional case where $(Z,X)(t') = (k',x')$. To be precise,

$$\mathrm{P}(\widetilde{Z}^{(t',k',x')}(t)=k \mid \widetilde{Z}^{(t',k',x')}(t')=k') = \mathrm{P}(Z(t)=k \mid (Z,X)(t')=(k',x')), \tag{4.2}$$

for all $t \geq t'$ and $k \in \mathcal{J}$, where t',k',x' is fixed. This can for example be utilised to calculate expectations of functions of $Z(t)$, by replacing with $\widetilde{Z}^{(t',k',x')}(t)$. For example, let g be some function, then

$$\begin{aligned}\mathrm{E}\left[g(Z(t))\,\middle|\,(Z,X)(t')=(k',x')\right] &= \mathrm{E}\left[g\left(\widetilde{Z}^{(t',k',x')}(t)\right)\,\middle|\,\widetilde{Z}^{(t',k',x')}(t')=k'\right]\\ &= \sum_k g(k)p_{k'k}(t;t',x').\end{aligned}$$

Theorem 4.4 yields the transition probabilities $p_{k'k}(t;t',x')$ from a simple differential equation, and thus, in the conditional case where $(Z,X)(t')=(k',x')$ and we know the associated forward transition rates, this will significantly reduce the complexity of the calculation.

The definition of the forward rates is useful beyond Theorem 4.4, as we see in Corollary 5.4 in Section 5 about life insurance.

4.1 Forward rates in the literature

Forward rates for doubly stochastic Markov chains have been discussed and treated in the literature since [14] introduced the concept of a forward mortality rate, and the forward mortality rate is the primary example of a forward transition rate. To the authors knowledge, a general definition for doubly stochastic Markov chains has not before been introduced. In this section, we briefly relate Definition 4.1 above to existing literature on forward rates.

The forward mortality rate is defined in the setup of a 2-state survival model with state space $\mathcal{J}=\{0,1\}$, where the only non-zero transition rate is from state 0 to 1, and this transition rate is simply denoted $\mu(t,X(t))$. State 0 corresponds to being alive, and state 1 corresponds to being dead, and we refer to $\mu(t,X(t))$ as the mortality rate. The forward mortality rate is defined as the function $g(t;t',x')$ that satisfies

$$p_{00}(t;t',x')=\mathrm{E}\left[e^{-\int_{t'}^{t}\mu(s,X(s))\mathrm{d}s}\,\middle|\,X(t')=x',Z(t')=0\right]=e^{-\int_{t'}^{t}g(s;t',x')\mathrm{d}s}. \tag{4.3}$$

In the literature, the measure used in (4.3) to define the forward mortality rate is a market consistent measure, and not necessarily the physical measure. In this paper we omit the discussion about which measure we operate with (and omit it from the notation), and simple note that for applications it is important to distinguish between relevant measures. For more on market consistent measures and market values, see [16].

The definition of the forward mortality rate (4.3) originates from [14], and has since been used in e.g. [6], [3] and [7]. The forward transition rate from Definition 4.1 is identical to the forward mortality rate in this setup. The quantity (4.3) can be differentiated on both sides to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}p_{00}(t;t',x')=-g(t;t',x')p_{00}(t;t',x').$$

By a comparison with (4.1), which simplifies in the simple 2-state setup, we can recognise the forward mortality rate as the forward transition rate in our setup, $g(t;t',x') = f_{01}(t;t',x')$. Further, using Lemma 4.3, we can represent the forward mortality rate as the expected mortality rate conditionally on being alive,

$$g(t;t',x') = \mathrm{E}\left[\mu(t,X(t))\,\middle|Z(t)=0, X(t')=x'\right].$$

There has been attempts at generalising the forward rates to more complex models than the 2-state survival model, for example if there are several dependent transition rates, or if the transition rate is dependent on some other stochastic process, e.g. a stochastic interest rate, see [15], [1] and [19]. Also, as is investigated in [5], the forward mortality rate as defined in (4.3) cannot be directly applied for more general models than the 2-state survival model. Definition 4.1 is more general, and while it specialises to the forward mortality rate in the 2-state model, it leads to more general forward rates in other models.

In this paper, only transition rates are considered, so any dependence on e.g. an interest rate is out of the scope of this treatment and postponed for further research. In [1], so-called *dependent forward rates* are defined. Consider the survival model with multiple causes of of death. The state space is $\mathcal{J} = \{0,\dots,J\}$, where state 0 is alive, and state $k > 0$ is interpreted as dead by cause k. Thus, the only non-zero transition rates are those out of state 0, thus all other states are absorbing. In this setup, the dependent forward rates are defined as the quantities $g_{0k}(t;t',x')$ satisfying

$$\begin{aligned} p_{00}(t;t',x') &= \mathrm{E}\left[e^{-\int_{t'}^{t}\sum_{\ell=1}^{J}\mu_{0\ell}(s,X(s))\,\mathrm{d}s}\,\middle|\, X(t')=x', Z(t')=0\right] \\ &= e^{-\int_{t'}^{t}\sum_{\ell=1}^{J} g_{0\ell}(s;t',x')\,\mathrm{d}s}, \end{aligned} \tag{4.4}$$

$$\begin{aligned} p_{0k}(t;t',x') &= \mathrm{E}\left[\int_{t'}^{t} e^{-\int_{t'}^{s}\sum_{\ell=1}^{J}\mu_{0\ell}(u,X(u))\,\mathrm{d}u}\mu_{0k}(s,X(s))\,\mathrm{d}s\,\middle|\, X(t')=x', Z(t')=0\right] \\ &= \int_{t'}^{t} e^{-\int_{t'}^{s}\sum_{\ell=1}^{J} g_{0\ell}(u;t',x')\,\mathrm{d}u} g_{0k}(s;t',x')\,\mathrm{d}s. \end{aligned} \tag{4.5}$$

Similar to the forward mortality rate, the dependent forward rates are defined by a replacement argument: It is the quantities that allow us to write the transition probabilities with the usual formulae as in the deterministic setup, and thereby get rid of the expectation. Analogously to above, one can differentiate $p_{00}(t;t',x')$ and $p_{0k}(t;t',x')$ and obtain differential equations for these transition probabilities. Then, one sees from (4.1) that the forward rates from Definition 4.1 specialises to the dependent forward rates in this setup. In particular, we can also represent the dependent forward rates as the conditional expectation of the transition rates, conditional on being alive,

$$g_{0k}(t;t',x') = \mathrm{E}\left[\mu_{0k}(t,X(t))\,\middle|Z(t)=0, X(t')=x'\right].$$

In [19], the concept of forward transition rates is examined and discussed. An overview is given, and related to the forward interest rate, from which the definition of the forward mortality rate is motivated, and a sceptical view is taken on the usefulness of forward rates. It is argued that in the general doubly stochastic setup, it does not seem possible to obtain forward transition rates that is defined in a meaningful way. It is the belief of the author, that to a certain extent, Definition 4.1 does exactly this, which is further supported by Lemma 4.3 and Theorem 4.4. It shall be noted, that the forward transition rates presented in this paper do not encompass dependence on e.g. the interest rate, which can be of interest within e.g. life insurance or credit risk.

4.2 Calculation of the forward rates

The forward differential equation presented in Theorem 4.4 does seem appealing for calculation of the transition probabilities, when compared with Corollary 3.2. However, to solve it one must first find the forward transition rates, which is not an easy task in practice. In certain simple models, it is however possible to find the forward transition rates. If an affine setup is considered, that is, where $\mathbf{X}$ is an affine process, the forward mortality rate from (4.3) can be found as a solution to certain ordinary differential equations, see e.g. [6] or [1]. Also, as is shown in [1], this result generalises to the case of the dependent forward rates from (4.4) and (4.5). The definition of the dependent forward rates indeed originate from the affine setup, where it is possible to calculate these.

In more complicated setups, it is not obvious how to calculate the forward transition rates and thereby gain a computational advantage from Theorem 4.4. Thus, Corollary 3.2 so far seems like the better choice for an actual calculation. It shall be noted, that if one solves (3.6) from Corollary 3.2, it is straightforward though to calculate the forward transition rates, since the distribution obtained from solving this is sufficient for calculating the forward rates by Lemma 4.3. This can be useful, for example for calculation of the cash flow in life insurance mathematics, using Corollary 5.4 below.

Even if the computational advantages of using the forward transition rates are not obvious in general models, the author believes that the forward transition rates are an interesting contribution towards understanding the structure of doubly stochastic Markov chains, as they provide a generalisation of the forward transition rates seen in the literature.

5 Life insurance cash flows

Let a finite state space $\mathcal{J}$ be given, where the states could be *alive*, *disabled*, *dead*, or similar. Then, let the doubly stochastic Markov chain $\mathbf{Z}$ from Section 3 describe the state of an insured in this state space. To each state k we associate a continuously paid payment rate $b_k(t)$, and to each transition we associate a payment, $b_{k\ell}(t)$. We assume both are continuous functions. The payments of the contract, accumulated until time t, is then described by the payment process $B(t)$, satisfying

$$\mathrm{d}B(t) = \sum_{k\in\mathcal{J}} 1_{(Z(t)=k)} b_k(t)\mathrm{d}t + \sum_{\substack{k,\ell\in\mathcal{J}\\ k\neq\ell}} b_{k\ell}(t)\mathrm{d}N_{k\ell}(t).$$

We assume that all payments occur before the finite time horizon T, i.e. that $b_k(s) = b_{k\ell}(s) = 0$ for all k, ℓ and $s \geq T$. This setup constitutes our modelling of the life insurance contract. Here we let b_k and $b_{k\ell}$ be deterministic functions, but we later discuss the straightforward extension to dependence on the underlying stochastic process $\mathbf{X}$.

We define the cash flow associated with this contract, as the expected payments.

Definition 5.1. *The* accumulated cash flow *at time t' conditional on $Z(t') = k'$ and $X(t') = x'$ is the function,*

$$A_{k'}(t; t', x') = \mathrm{E}\left[B(t) - B(t') \middle| Z(t') = k', X(t') = x' \right].$$

Furthermore, if $A_{k'}(t; t', x')$ has a density with respect to the Lebesgue measure,

$$A_{k'}(\mathrm{d}t; t', x') = a_{k'}(t; t', x')\,\mathrm{d}t,$$

we refer to $a_{k'}(t; t', x')$ as the cash flow.

Remark 5.2. In the traditional life insurance setup, single payments can also happen at deterministic time points. In that case, one allows the cash flow to have a density with respect to a mixture between the Lebesgue measure and a counting measure. An extension to this more general case is straightforward, but complicates the notation and is thus omitted from the present article. ◇

One can now show the following result; a proof in the semi-Markov setup is given in [2], and the proof in the present setup is essentially identical.

Proposition 5.3. *The cash flow exists and is given by,*

$$a_{k'}(t; t', x') = \sum_{k\in\mathcal{J}} \int_{\mathcal{X}} P(t, k, \mathrm{d}x; t', k', x') \Bigg(b_k(t) + \sum_{\substack{\ell\in\mathcal{J}\\ \ell\neq k}} \mu_{k\ell}(t, x) b_{k\ell}(t) \Bigg). \tag{5.1}$$

The cash flow can also be represented as a function of the transition probabilities for $\mathbf{Z}$ and the forward transition rates.

Corollary 5.4. *The cash flow has representation*

$$a_{k'}(t;t',x') = \sum_{k\in\mathcal{J}} p_{k'k}(t;t',x')\left(b_k(t) + \sum_{\substack{\ell\in\mathcal{J}\\ \ell\neq k}} f_{k\ell}(t;t',k',x')b_{k\ell}(t)\right). \tag{5.2}$$

Proof. We have that $P(t,k,\mathrm{d}x;t',k',x')\mathrm{d}\tau_{\mathcal{J}}(k)$ is the probability measure of $(Z,X)(t)$ conditional on $(Z,X)(t') = (k',x')$, and that $p_{k'k}(t;t',x')\mathrm{d}\tau_{\mathcal{J}}(k)$ is the corresponding marginal probability measure of $Z(t)$. Thus,

$$\frac{P(t,k,\mathrm{d}x;t',k',x')}{p_{k'k}(t;t',x')}$$

is the conditional probability measure of $X(t)$ given both $Z(t) = k$ and $(Z,X)(t') = (k',x')$. Using this, we find

$$\begin{aligned}
&\int_{\mathcal{X}} \mu_{kl}(t,x)P(t,k,\mathrm{d}x;t',k',x') \\
&= p_{k'k}(t;t',x')\int_{\mathcal{X}} \mu_{kl}(t,x)\frac{P(t,k,\mathrm{d}x;t',k',x')}{p_{k'k}(t;t',x')} \\
&= p_{k'k}(t;t',x')\,\mathrm{E}\left[\mu_{kl}(t,X(t))\,\middle|\,Z(t)=k,(Z,X)(t')=(k',x')\right] \\
&= p_{k'k}(t;t',x')f_{kl}(t;t',k',x'),
\end{aligned}$$

where we in the last equality applied Lemma 4.3. Insertion into Proposition 5.3 yields the result. ☐

In the Corollary, the integral from formula (5.1) disappears, and we recognise formula (5.2) as the cash flow formula from the classic setup with deterministic transition rates. Indeed, if the transition rates are deterministic, $\mu_{k\ell}(t,x)$ does not depend on x and the integral disappears from (5.1), while also the transition rates equal the forward transition rates. Thus, in this case (5.1) and (5.2) equal.

Given some continuously compounded interest rate $r(t)$, which we assume to be deterministic, we can calculate the expected present value at time t, conditional on $Z(t) = k$. This is defined as,

$$V_k(t,x) = \mathrm{E}\left[\int_t^T e^{-\int_t^s r(w)\mathrm{d}w}\mathrm{d}B(s)\,\middle|\, Z(t)=k, X(t)=x\right].$$

One can show that the expected present value is simply the sum of the discounted cash flow, which we state in the following proposition.

Proposition 5.5. *If r is deterministic, the expected present value at time t, conditional on $Z(t) = k$ and $X(t) = x$ is given as*

$$V_k(t,x) = \int_t^T e^{-\int_t^s r(w)\mathrm{d}w} a_k(s;t,x)\mathrm{d}s.$$

The main objective of the actuary is to calculate the expected present value, and this can be done e.g. by calculating the cash flow first, and then discounting it, as described above. Another possibility is to calculate it directly, since one can show that $V_k(t,x)$ is the solution of a backward PDE, similar in structure to the backward PDE for the transition probability, (3.8), see e.g. [7]. However, one is often interested in the cash flow, since it is convenient for an analysis of the interest rate risk in practice. For example, if one wants to calculate the expected present value with different interest rates, it is advantageous to first calculate the cash flows, which are independent of the interest rate. In particular, if a large portfolio of insurance contracts is considered, the cash flows can be accumulated to a single cash flow for the portfolio, which are easily discounted. For calculation of the cash flow, one needs the transition probabilities, which can be calculated with Corollary 3.2.

Remark 5.6. Mathematically it is straightforward to extend the current setup such that the payment functions b_k and $b_{k\ell}$ may depend on the value of the underlying process $\mathbf{X}$, and we can write $b_k(t,x)$ and $b_{k\ell}(t,x)$. In that case, the cash flow (5.1) would be

$$a_{k'}(t;t',x') = \sum_{k\in\mathcal{J}} \int_{\mathcal{X}} P(t,k,\mathrm{d}x;t',k',x') \left(b_k(t,x) + \sum_{\substack{\ell\in\mathcal{J}\\ \ell\neq k}} \mu_{k\ell}(t,x) b_{k\ell}(t,x) \right).$$

Note that it is not possible write the cash flow as a function of the forward transition rates as in Corollary 5.4. ◇

Remark 5.7. In this section, we assumed a deterministic interest rate $r(t)$, however a stochastic interest rate is easily handled if it is independent of the underlying stochastic process $\mathbf{X}$ and $\mathbf{Z}$. For more on stochastic interest rates in life insurance, see e.g. [16] and references therein. ◇

6 Semi-Markov models in life insurance

If we extend the classic life insurance setup, where the transition rates are deterministic, to allow the transition rates and the payment functions to depend on the time spent in the current state, we have introduced duration dependence in the setup. In this case, $\mathbf{Z}$ is a semi-Markov process, and this class of models are popular in e.g. life insurance. An example of such a model is a disability model with recovery, where the recovery rate and the mortality as disabled might decrease as a function of the time spent as disabled. For a treatment of this model, see [2] and references therein.

We construct the semi-Markov model and show that the semi-Markov process is a special case of the process in Section 2. This can be used to present Theorem 2.1 for the semi-Markov process, and in this case it becomes an integro-differential equation. This result is also presented in [2], but the proof of the result is different.

Let $\mathbf{Z}$ be a stochastic process on a finite state space $\mathcal{J}$. Define $\mathbf{U}$ as the duration in the current state,

$$U(t) = \sup\left\{s \in [0,t] \mid Z(w) = Z(t), w \in [t-s,t]\right\}.$$

We assume that $(\mathbf{Z}, \mathbf{U})$ is a Markov process. The process $(\mathbf{Z}, \mathbf{U})$ may jump, and the pure jump part can be written as

$$J(t) = \sum_{s \le t} \begin{bmatrix} \Delta Z(s) \\ \Delta U(s) \end{bmatrix}.$$

We denote the jump measure of $J(t)$, as defined in (2.2), by $\tilde{\mu}(\mathrm{d}(\ell, v); t, (k,u))$, and it has a density with respect to a counting measure,

$$\tilde{\mu}(\mathrm{d}(\ell, v); t, (k,u)) = \mu_{k\ell}(t,u) \cdot \mathrm{d}\left(\tau_{\mathcal{J}\setminus\{k\}}(\ell) \otimes \tau_{\{0\}}(v)\right).$$

Here, $J(t)$ is the jump sizes, and $\tilde{\mu}((\ell, v); t, (k,u))$ is interpreted as the instantaneous probability that if we are in state k with duration u, we make a jump to state ℓ and have duration v. By the definition of $\mathbf{U}$ we always jump to duration 0, so if $v \neq 0$ the density is zero. We interpret $\mu_{k\ell}(t,u)$ as the transition rate of $\mathbf{Z}$, and it can be shown to satisfy the relation

$$\mu_{k\ell}(t,u) = \lim_{\delta \searrow 0} \frac{1}{\delta} P\left(Z(t+\delta) = \ell | Z(t) = k, U(t) = u\right),$$

for $k \neq \ell$. We interpret this as the instantaneous probability of a jump of $\mathbf{Z}$ from state k to state ℓ, if we are at time t with duration u.

With $N_{k\ell}(t)$ given by (3.1), we can characterise $U(t)$ as

$$\mathrm{d}U(t) = 1\mathrm{d}t - U(t-) \sum_{\substack{k,\ell \in \mathcal{J} \\ k \neq \ell}} \mathrm{d}N_{k\ell}(t).$$

The first term on the right hand side is the constant increase of $U(t)$ with slope 1, and the second term states that whenever $\mathbf{Z}$ jumps, $\mathbf{U}$ jumps to zero. We remark that the state space of $U(s)$ is $[0,s]$, and, conditional on $U(t) = u$, the state space for $U(s), s \geq t$ is $[0, u+s-t]$.

We define the transition probability

$$P(t,k,u;t',k',u') = \mathrm{P}(Z(t)=k, U(t)\le u \mid Z(t')=k', U(t')=u').$$

We specialise Theorem 2.1 to this semi-Markov setup. Because of the structure of the process, in particular the fact that the duration process increases identically as time almost everywhere and possess no diffusion term, we can write it as an integro-differential equation (IDE). We note, that since $\frac{\partial}{\partial u}P(t,k,u;t',k',u')$ does not exist for all u, the result is not a direct special case of Theorem 2.1.

Theorem 6.1. *The transition probability $P(t,k,u;t',k',u')$, for $t\ge t'$, $k\in\mathcal{J}$ and $u\in[0,u'+t-t']$ satisfies the forward IDE*

$$\begin{aligned}\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial u}\right)P(t,k,u;t',k',u') = &\sum_{\substack{\ell\in\mathcal{J}\\ \ell\neq k}}\int_0^{u'+t-t'} \mu_{\ell k}(t,v)P(t,\ell,\mathrm{d}v;t',k',u') \\ &-\int_0^u \sum_{\substack{\ell\in\mathcal{J}\\ \ell\neq k}}\mu_{k\ell}(t,v)P(t,k,\mathrm{d}v;t',k',u'),\end{aligned} \tag{6.1}$$

subject to the boundary conditions $P(t',k,u;t',k',u') = 1_{(k=k')}1_{(u'\le u)}$ and, for $t>t'$, $P(t,k,0;t',k',u')=0$.

The differential equation can be interpreted both as a partial integro-differential equation, and as an ordinary integro-differential equation. The latter interpretation makes it easy to solve in practice, and for details on this see [2], where an algorithm is explained. The theorem is proven in [2], however, we give an outline of how to prove it with the tools presented in Section 2. Theorem 2.1 yields the IDE on the interior of the domain, so what is left is to make sure the IDE also holds on the boundary.

For the proof and subsequent results, we introduce the following notation,

$$\begin{aligned}p_{\overline{kk}}(t;t',u') &= \mathrm{P}(Z(t)=k, U(t)=u'+t-t' \mid Z(t')=k, U(t')=u') \\ &= \exp\left\{-\int_0^{t-t'}\sum_{\substack{\ell\in\mathcal{J}\\ \ell\neq k}}\mu_{k\ell}(t'+w,u'+w)\mathrm{d}w\right\}.\end{aligned} \tag{6.2}$$

This quantity is the probability of staying in state k from time t' with duration u' until time t. It is well known that it has the above expression, see e.g. [11].

Proof. (Outline) The first boundary condition follows directly from the definition of the transition probability. Since, for any t, the probability that a jump occurs exactly at time t can be shown to equal 0, the probability that the duration is equal to zero is zero, for all t. This yields the second boundary condition.

We first claim that one can show, that on the open set

$$\{(t,u) \mid t \in (t',T), u \in (0, u'+t-t')\},$$

the partial derivatives $\frac{\partial}{\partial t}P(t,k,u;t',k',u')$ and $\frac{\partial}{\partial u}P(t,k,u;t',k',u')$ exist. This holds, since probability mass on that set originates from jumps. Thus, from Theorem 2.1, the result holds on this set.

Since $P(t,k,u;t',k',u')$ is right-continuous in u by definition, the result holds for $u=0$. For $u = u'+t-t'$, note that

$$\begin{aligned} P(t,k,u'+t-t';t',k',u') = {} & P(t,k,u'+t-t';t',k',u') - P(t,k,u'+t-t'-\delta;t',k',u') \\ & + P(t,k,u'+t-t'-\delta;t',k',u'), \end{aligned}$$

and let $\delta \searrow 0$ to obtain

$$\begin{aligned} P(t,k,u'+t-t';t',k',u') = {} & \mathrm{P}(Z(t)=k, U(t)=u'+t-t' \mid Z(t')=k', U(t')=u') \\ & + P(t,k,(u'+t-t')-;t',k',u'). \end{aligned} \tag{6.3}$$

The second term on the right hand side of (6.3) satisfies (6.1). The first term on the right hand side is, for $k=k'$, the probability of staying in state k' from time t' to t, and from (6.2) we know that this probability is of the form

$$\mathrm{P}(Z(t)=k, U(t)=u'+t-t' \mid Z(t')=k', U(t')=u') = 1_{(k=k')} p_{\overline{kk}}(t;t',u').$$

Applying $\frac{\partial}{\partial t} + \frac{\partial}{\partial u}$ to both sides, we find that this functions satisfies a linear partial differential equation of the form (6.1).

Since both terms on the right hand side of (6.3) satisfy (6.1), which is linear, we use the fact that if two functions satisfy the same linear IDE, each with their own boundary condition, the sum of of these functions satisfy the same linear IDE, with the boundary condition being the sum of the two boundary conditions. From this we conclude that the whole of (6.3) satisfy (6.1). Thus, the result holds for all $u \in [0, u'+t-t']$ and $t \geq t'$. $\square$

We can partly solve the forward IDE, and obtain the following integral equation, which is essentially a differentiated version of equation (4.7) in [11].

Proposition 6.2. *The transition density $p(t,k,u;t',k',u')$, for $k \in \mathcal{J}, u \in [0, u'+t-t']$ exists with respect to a mixture of the Lebesgue measure and a point measure. For fixed t,*

$$P(t,A,B;t',k',u') = \sum_{k \in A} \int_B p(t,k,u;t',k',u') \left(\mathrm{d}u + \mathrm{d}\tau_{\{u'+t-t'\}}(u)\right).$$

The density is given by

- *for $t - t' > u$, the integral equation*

$$p(t, k, u; t', k', u') = \sum_{\substack{\ell \in \mathcal{J} \\ \ell \neq k}} \int_0^{u'+t-u-t'} p(t-u, \ell, v; t', k', u') \mu_{\ell k}(t-u, v) \times p_{\overline{kk}}(t; t-u, 0) \left(\mathrm{d}v + \mathrm{d}\tau_{\{u'+t-u-t'\}}(v) \right),$$

- *else,*

$$p(t, k, u; t', k', u') = 1_{(k=k')} 1_{(t-t'=u-u')} p_{\overline{kk}}(t, t', u').$$

Proof. (Outline) For $t - t' > u$, a jump must have occurred between time t' and time t, and we argued in the proof of Theorem 6.1 that a density with respect to the Lebesgue measure exists. That the proposed density satisfies the IDE (6.1) can be seen by applying $\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial u}\right)$ to

$$P(t, k, u; t', k', u') = \int_0^u p(t, k, v; t', k', u') \left(\mathrm{d}v + \mathrm{d}\tau_{\{u'+t-t'\}}(v) \right),$$

and where the proposed density is inserted.

For the second part where $t - t' \leq u$, a jump cannot have occurred after time t', and we must have stayed in the initial state. Then the result is known from (6.2). □

We interpret the terms of the density. As was argued in the proof, the last part is the probability mass originating from the initial state, and is simply the probability of having made no jumps. The probability mass in the first part is all jumps into state k happening at time $t - u$. Inside the integral, we see the probability of being in a state ℓ at time $t - u$ with a duration v, multiplied with the probability of a jump to state k, which is again multiplied with the probability of then staying in state k from time $t - u$ to time t. Then we simply sum over all states ℓ and all durations v.

Acknowledgements

I would like to thank Martin Jacobsen, Rasmus Hedegaard Kristiansen, Thomas Møller and Mogens Steffensen for fruitful discussions during this work. I also thank an anonymous referee for suggestions and comments that improved the paper.

References

[1] Kristian Buchardt. Dependent interest and transition rates in life insurance. *Insurance: Mathematics & Economics*, 55:167–179, 2014.

[2] Kristian Buchardt, Thomas Møller, and Kristian Bjerre Schmidt. Cash flows and policyholder behaviour in the semi-Markov life insurance setup. To appear in Scandinavian Actuarial Journal, 2014.

[3] Andrew JG Cairns, David Blake, and Kevin Dowd. Pricing death: Frameworks for the valuation and securitization of mortality risk. *Astin Bulletin*, 36(1):79, 2006.

[4] Marcus C. Christiansen. Multistate models in health insurance. *Advances in Statistical Analysis*, 96(2):155–186, 2012.

[5] Marcus C. Christiansen and Andreas Niemeyer. On the forward rate concept in multi-state life insurance. *Finance and Stochastics*, 19(2):295–327, 2015.

[6] Mikkel Dahl. Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance: Mathematics & Economics*, 35 (1):113–136, 2004.

[7] Mikkel Dahl and Thomas Møller. Valuation and hedging of life insurance liabilities with systematic mortality risk. *Insurance: Mathematics & Economics*, 39 (2):193–217, 2006.

[8] Mark H. A. Davis and Michel H. Vellekoop. Permanent health insurance: A case study in piecewise-deterministic markov modelling. *Mitteilungen der Schweiz. Vereinigung der Versicherungsmathematiker*, 2:177–211, 1995.

[9] Crispin Gardiner. *Stochastic methods: A handbook for the natural and social sciences*. Springer Series in Synergetics. Springer, 4th edition, 2009.

[10] Marko Helwich. Durational effects and non-smooth semi-Markov models in life insurance. `urn:nbn:de:gbv:28-diss2008-0056-4`, 2008. Doctoral dissertation, University of Rostock.

[11] Jan M. Hoem. Inhomogeneous semi-Markov processes, select actuarial tables, and duration-dependence in demography. *T. N. E. Greville (ed.), Population Dynamics*, pages 251–296, 1972.

[12] Michael Koller. *Stochastic models in life insurance*. European Actuarial Academy Series. Springer, 2012.

[13] David Lando. On Cox processes and credit risky securities. *Review of Derivatives Research*, 2:99–120, 1998.

[14] Moshe A. Milevsky and David Promislow. Mortality Derivatives and the Option to Annuitise. *Insurance: Mathematics & Economics*, 29 (3):299–318, 2001.

[15] Kristian R. Miltersen and Svein-Arne Persson. Is Mortality Dead? Stochastic Forward Force of Mortality Rate Determined by No Arbitrage. (Working Paper), Norwegian School of Economics and Business Administration, 2005.

[16] Thomas Møller and Mogens Steffensen. *Market-valuation methods in life and pension insurance*. International Series on Actuarial Science. Cambridge University Press, 2007.

[17] Christian Max Møller. A stochastic version of Thiele's differential equation. *Scandinavian Actuarial Journal*, 1:1–16, 1993.

[18] Ragnar Norberg. Reserves in life and pension insurance. *Scandinavian Actuarial Journal*, 1991:1–22, 1991.

[19] Ragnar Norberg. Forward mortality and other vital rates - Are they the way forward? *Insurance: Mathematics & Economics*, 47 (2):105–112, 2010.